

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1369	54.0	488	4	Q9NVB5
2	717	30.5	339	4	Q9UHE8
3	717	30.5	373	4	Q95034
4	577	24.5	264	4	Q9R605
5	156.5	6.7	239	2	Q9RKR6
6	147	6.3	213	1	Q29059
7	142	6.0	212	1	Q29370
8	142	6.0	224	1	Q59661
9	136.5	5.8	217	2	Q9JN78
10	119.5	5.1	216	2	Q50637
11	116.5	5.0	242	2	Q9PCPL
12	114	4.8	207	2	Q9RIU0
13	112.5	4.8	207	2	Q9RII8
14	112	4.8	408	5	Q9V8X8
15	111.5	4.7	1228	3	Q12754
16	109	4.6	219	2	Q9RII2
17	109	4.6	320	12	Q9YVM4
18	109	4.6	1301	10	Q23530
19	105.5	4.5	348	8	Q79242

ALIGNMENTS

RESULT 1

PREL.TMTNARY: PRT. 488 AA

AC	O9NVB5:	
DT	01-OCT-2000	(TREMBLrel. 15, Created)
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)
DE	CDNA FLJ10829 FIS, CLONE NT2RP4001138.	

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N. A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Tshii S., Kawai Y., Saito K., Yamamoto I., Wakamats

RA · Nakamura Y., Nagahar1 K., Masuho Y., Sasaki N.;

RT "NEDO human cDNA sequencing project." :

Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

SQ SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;

```
Query Match
54.0%; Score 1269; DB 4; Length 488;
```

Best Local Similarity 54.3%; Pred. No. 3.4e-89;

Matches 241; Conservative 83; Mismatches 108

3 SI5MMGSPKSLSETCLPNCINGIKDARKVTGVIGSGDFAKSLITRLTRCGYHV

Age Group	Total	Male	Female	Male	Female
18-24	15%	10%	20%	5%	15%
25-34	25%	20%	30%	15%	25%
35-44	30%	25%	35%	20%	30%
45-54	20%	15%	25%	10%	20%
55-64	15%	10%	20%	5%	15%
65-74	10%	5%	15%	3%	10%
75+	5%	3%	10%	2%	5%

Db 11 SLHLVDSDDLAK--VPD-----EAPK--VGILGSGDFARSLATRLVCGSGFKVV

09 63 PKFSEFEPHVVDVTHHEDAI TKNTTFVA I HRFHYTSI WDIBH I VCKI I IDVS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 61 PKRTALYPAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLDQLAGKILVDVS

131 PTNOVDESNAPVIASSTEDOSTTUVCEPNN/CANWATOICBPKACACDIOUYTCCNNTC

QY IZI KINQIFESNHELASLEF PDSLEIVAGFN V 38WHLQDGPDRQV IIC3NNI QAR

Db 121 EHLQHRESNAEYLASLFPTCTVVKA FNVISAWTLQAGPRDGNROVPICGDQPEAK

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 181 LARQLNFIPDLGSLSSAREIENLPRLRFTLWRGPVVAISLATFFFLYSFVRD

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities related to the business. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various methods and techniques used to collect and analyze data. It includes a detailed description of the experimental setup and the procedures followed during the data collection process.

3. The third part of the document presents the results of the study, showing the trends and patterns observed in the data. It includes a comparison of the findings with previous research and a discussion of the implications of the results.

4. The fourth part of the document discusses the limitations of the study and suggests areas for future research. It highlights the need for further investigation into the factors that influence the outcomes of the study.

5. The fifth part of the document provides a conclusion and summarizes the key findings of the study. It reiterates the importance of accurate record-keeping and the need for ongoing research in this field.

Db 181 MALAMFMPVDMGSLASAWVEAMPLRLPAWKVPTLLALGLFVCFYAYNEVRDVLQPVV 240
 QY 241 RNOQSDFYKPIETVNTKPLTPIVAITLLSLVYLAGLLAAAYOLYGTGYRRPWPWLEWLO 300
 Db 241 QESONKFFKLPVSVNTPLPCVAVVLLSVLPGVLAALQLRGTKYQRPFDWLDHWO 300
 QY 301 CRKQGLLSFFAMVHVAYSICLPMRSEYRLFENMAYQOVHANIENSWNEEVRRIEMY 360
 Db 301 HRKQIGLLSFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRMEIY 360
 QY 361 ISFGINSGLLSLNAVSPSVSNALNWRPESFQSLGIVYVALLISTFHVLIYKKAPE 420
 Db 361 LSLGVLAGLTLSSLAVTSPLSIANSANLWREESFVQSSIGFVALVLSLTLTLYGWTAFE 420
 QY 421 EBYVRFYTPPNFVALVLPISVIL 444
 Db 421 ESRKFLPPTFTLLVPCVIL 444

RESULT 2

ID Q9UHE8 PRELIMINARY; PRT; 339 AA.
 AC Q9UHE8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE SIX TRANSMEMBRANE EPITHELIAL ANTIGEN OF PROSTATE.
 GN STEAP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,
 Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,
 Jakobovits A., Sadrir D.C., Afar D.E.H.;
 RT "STEAP: A prostate-specific cell-surface antigen highly expressed in
 human prostate tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
 DR EMBL: AF186249; AAL17479.1;
 KW Transmembrane.
 SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 4; Length 339;
 Best Local Similarity 54.9%; Pred. No. 3.7e-47;
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVVAISLATFFFLYSEVRDVIHPYARNQSDFYKPIETVNTKPLTPIVAITLL 267
 Db 67 LFPQWHLPIKIAAIIASITFLYTLREVIHPLATSHOQYFYKPIILVINKVLPVVSITLL 126
 QY 268 SLVYLAGLLAAAYOLYGTGYRRPWPWLEWLOCRKQGLLSFFAMVHVAYSICLPMR 327
 Db 127 ALVYLPGVIAAVOLHNGTKYKKEPPLDKWMLTRKQGLLSFFAVLHAIYLSYPMR 186
 QY 328 SERYLEFNMAVQOVHANIENSWNEEVRRIEMYISFGIMSLGLLSLAVTSIPSVSNALN 387
 Db 187 SYRKLLNWAYQOVQONKEDAWIHDVWMEIYVSLGIVGLAILALLAVTSIPSVSOLT 246
 QY 388 WREFSFQSTLGYVALLISTFHVLIYKKAPEEYRYFTPPNFVALVLPISVIL 444
 Db 247 WREFHYIQSKIGIVSLGLTTHALIFAWNKWIDIKQFWVTPPTFMIAVFLPIVILI 303

RESULT 3

ID Q95034 PRELIMINARY; PRT; 373 AA.
 AC Q95034
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-AUG-1999 (TrEMBLrel. 11, Last annotation update)

DE WUGSC:H_RG041D11.1 PROTEIN (WUGSC:H_DJ1121E10.1 PROTEIN) (FRAGMENT)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
 RT "The sequence of Homo sapiens BAC clone RG041D11."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 25-373 FROM N.A.
 RX MEDLINE-99063792; PubMed-9847074;
 RA Sulston J.E., Waterston R.;
 RT "Toward a complete human genome sequence."
 RL Genome Res. 8:1097-1108(1998).
 RN [5]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Kraemer J., Miller N., Ozersky P., Walker C.;
 RT "The sequence of Homo sapiens PAC clone DJ1121E10."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Waterston R.H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Waterston R.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005053; AAC79150.1;
 DR EMBL: AC004969; AAD15620.1;
 FT NON-TER 1;
 SQ SEQUENCE 373 AA; 43574 MW; B8BE3C86F903F775 CRC64;

Query Match 30.5%; Score 717; DB 4; Length 373;
 Best Local Similarity 54.9%; Pred. No. 4.2e-47;
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVVAISLATFFFLYSEVRDVIHPYARNQSDFYKPIETVNTKPLTPIVAITLL 267
 Db 101 LFPQWHLPIKIAAIIASITFLYTLREVIHPLATSHOQYFYKPIILVINKVLPVVSITLL 160
 QY 268 SLVYLAGLLAAAYOLYGTGYRRPWPWLEWLOCRKQGLLSFFAMVHVAYSICLPMR 327
 Db 161 ALVYLPGVIAAVOLHNGTKYKKEPPLDKWMLTRKQGLLSFFAVLHAIYLSYPMR 220
 QY 328 SERYLEFNMAVQOVHANIENSWNEEVRRIEMYISFGIMSLGLLSLAVTSIPSVSNALN 387
 Db 221 SYRKLLNWAYQOVQONKEDAWIHDVWMEIYVSLGIVGLAILALLAVTSIPSVSOLT 280
 QY 388 WREFSFQSTLGYVALLISTFHVLIYKKAPEEYRYFTPPNFVALVLPISVIL 444
 Db 281 WREFHYIQSKIGIVSLGLTTHALIFAWNKWIDIKQFWVTPPTFMIAVFLPIVILI 337

RESULT 4

ID Q9Y605 PRELIMINARY; PRT; 264 AA.
 AC Q9Y605
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE WUGSC:H_RG087E15.1 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
 OX NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9063792; PubMed=9847074;
 RA Sulston J.E., Waterston R.,
 RT "toward a complete human genome sequence.";
 RL Genome Res. 8:1097-1108(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strong C., Layman D., Graves T., Stromatt C.,
 RT "The sequence of Homo sapiens BAC clone RG087E15";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.,
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.,
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Waterston R.,
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005061; AAD43182.1;
 FT NON_TER 1 264
 FT SEQUENCE 264 AA; 31203 MW; 35C9483003557E72 CRC64;

Query Match 24.5%; Score 577; DB 4; Length 264;
 Best Local Similarity 56.4%; Pred. No. 1.4e-36;
 Matches 106; Conservative 36; Mismatches 46; Indels 0; Gaps 0;

QY 208 LFTLRGPPVVAISLATEFFLYSFVRDINHPYARNOQSDFYKIPIETVYNKPLPIVAIPL 267
 DB 77 LFPQHLPIKIAANVASLTFLYTLREVINHPLATSHQYFYKIPILVINKVLPVVSITLL 136
 QY 268 SLVYLGLAAAYQLYGYCYKRRPPPLWETLQCRKQLGLLSFFFAWVAVSYCLPMRR 327
 DB 137 ALVYLPVIAIVQVNTKTKKFFPHNDKWLTKRQGLLSLFAVLAHYLSYAMRR 196
 QY 328 SERYLFLAMAYQQVHANTNSWNEEVRRIEYISFGIMSLGLSLAVTSIPSYSNALN 387
 DB 197 SYRYKLLNWAYQQVQONKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSYSDSLT 256
 QY 388 WREPSFIQ 395
 DB 257 WREPHYIQ 264

RESULT 5
 QY 99RKR6 PRELIMINARY; PRT; 239 AA.
 AC 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE PUTATIVE OXIDOREDUCTASE.
 GN SCC75A.08C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_taxid=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=A3(2);
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redenbach M., Kleser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL133220; CAB61708.1;
 SQ SEQUENCE 239 AA; 24946 MW; 72B3F159A363EDE6 CRC64;

Query Match 6.7%; Score 156.5; DB 2; Length 239;
 Best Local Similarity 27.9%; Pred. No. 0.00018;
 Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 IKDARKVTGVT-GSGDFAKSLTIRLCRGYHVVTGSR---NPKFASEFFPHVDVTHHE 80
 DB 23 LPDVSGLVGVGGTPOCKGLAYRLAKAGQKIVIVGSRAAARAAAABEIGHGVEGADNA 82
 QY 81 DALTKTNIIFVAIHRE-HYTSIMDLRLHLLVCKILIDVSNMRLNQ-----YPE--SNAE 131
 DB 83 ETARSDVVIIVAVPDWGHGKTLSELRSLGKLVVDCVNLGDFGRKAYALKPEGSAAE 142
 QY 132 YLASLFPDSLVKGFNVVSAMALQ-----LGPDKASRQVYCSNNICARQOV 178
 DB 143 QAAALLPDSRVAAPFHLNSAVLLQDPEIDETDVMVIGEERADVEI-----VQA---- 192

RESULT 6
 QY 029059 PRELIMINARY; PRT; 213 AA.
 AC 029059;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DE 01-AUG-1998 (Tremblrel. 07, Last annotation update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN AF1209.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus
 OX NCBI_taxid=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001021; AAB90038.1;
 RT TIGR; AF1209;
 KW Hypothetical protein.
 SQ SEQUENCE 213 AA; 22994 MW; D0D7FD40A6A6AB05 CRC64;

QY	201	IENLPLRLFTLWR	213
	:	: :	:
Db	188	VESLTPLILNMR	200
RESULT	8		
ID	O59661	PRELIMINARY;	PRT; 224 AA.
AC	O59661;		
DT	01-AUG-1998 (TREMBLrel. 07, Created)		
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)		
DE	F420-DEPENDENT NADP REDUCTASE.		
OS	Methanobacterium thermoautotrophicum.		
OC	Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;		
OX	Methanobacterium.		
OCX	NCBI_TaxID=2166;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-WARBURG;		
RA	Berk H., Thauer R.K.;		
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: Y17210; CAA76687.1;		
SQ	SEQUENCE 224 AA; 23448 MW; 8B3514EC14BCB0CF CRC64;		
Query Match	6.0%; Score 142; DB i; Length 224;		
Best Local Similarity	29.2%; Pred. No. 0.0022;		
Matches	59; Conservative 29; Mismatches 84; Indels 30; Gap		
QY	37	GSGDFAKSLTIRLCGYHIVIGSNPKFASEFFPHVVDVTHHEDALTK--TN-----	8
Db	8	GTGDDGLGLALRIAGLEVIIIGSDAEKAVSAQKVLEIAERDLDLVKGATNAEAAPEA	6
QY	88	---IIFVAIHREHYTSMDLRHLVLKILIDV-----SNNNRINQYPESNAEYLA	1
Db	68	EVALITPTLQAQMAT-LGSVKEAIKGVLI DATVPIDSCLGGSVARYILDWGSAAERAA	1
QY	135	SLFPPD--SLIVKGFNVVSAMALQ--LGPKDASRQYVICNNIQARQQVIELARLNFT-P	1
Db	127	REFLEDQGTVAAFANISASALLDITGPVDC--DCLIASDHRLDALDLASELAEKIDGVRA	1
QY	190	IDGLSSLSARETENPLRLFTL	211
Db	185	IDCGGLENAVRIKITPLLINL	206
RESULT	9		
ID	O9JN78	PRELIMINARY;	PRT; 217 AA.
AC	O9JN78		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DE	PUTATIVE OXIDOREDUCTASE.		
GN	MMYO.		
OS	Streptomyces coelicolor.		
OC	Plasmid pSCPI.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OCX	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.		
OX	NCBI_TaxID=1902;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-A3(2);		
RA	Bruton C.J.; Wietzorrek A., Hartley N., Woodburn L., Chater K.F.;		
RL	"Genes involved in methylenomycin biosynthesis from plasmid SCPI of		
RT	Streptomyces coelicolor A3(2)."		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-A3(2);		
RX	MEDLINE=88112873; PubMed=2828187;		

RA MEDLINE=20190606; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.57-
RA Brill J.F., Adayani A., An H.-J., Andrews-Fpankoch C., Baldwin D.
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Duran-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush E., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris B., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003795; AAC57539.1;
 DR FLTBASE; FBgn0034456; CG15121.
 SQ SEQUENCE 408 AA; 47829 MW; 364769862D735B33 CRC64;

Query Match 4.8%; Score 112; DB 5; Length 408;
 Best Local Similarity 19.8%; Pred. No. 0.94; 137; Indels 112; Gaps 17;
 Matches 76; Conservative 59; Mismatches 137;
 Qy 132 YLASLPDLSL-----IVKGFVNSAWALQGLPKDASROVYICSNINQAOQVI 179
 Db 51 YHYOLFDSLSLPKKSVEQDIISGKYNLSLHGVIIRPEET-----SDFNATQHSY 103
 Qy 180 ELAROLNPIPIDLSLSARELENIPLRLFTLWRGP-----VVVALSLATFFLY 229
 Db 104 PLELTNCVMVPLA-----PELPKWMYVW--PLGKYIWTCLFLGTGFYVALLRYVH 153
 Qy 230 -----SFVROVIHPYARNOQSDFKIPIEIVNKTLPITVAITLSIVYLAGLIAAY 280
 Db 154 WREPONATRSYTRNVHLHALLMFSAANNMVKL--KHASIRVITFYTLTYFGFILTNY 211
 Qy 281 QLYYGTKYRRPPPL--ETWLOCRKQZGLLSFFFAFVHVAYSLCPMRSSRYFLPNA 337
 Db 212 HLSHWTAEDMKPVFLRPDTHSD-----LHSLRIVIHDSLLEELRWLPV- 257
 Qy 338 YOOVHAN-----IENSW---NEEVWRIEMV-----ISFGIMSLGLLSLAVTSIP 380
 Db 258 YOALLASPSRSYAVVTQDAMLFNRQOKVLIQIYFHLKVKVCFG---GLFNALPMASNA 313
 Qy 381 SVSNALN-----WREFSF-IQSTGLGVALLISTFHVLIYGNKRAPEEYR 425
 Db 314 SPADSLNKFIINWQAGLWNYEELAFYAEQYAKVFLDTPY-----EPLNLE 364
 Qy 426 FYTPNFVILVPSIVIDLQLO 449
 Db 365 FTTAWIVLSAGIPISSIAFLCLEL 388
 RESULT 15
 Q12754 PRELIMINARY; PRT; 1228 AA.
 ID Q12754
 AC Q12754;
 DT 01-NOV-1996 (TrEMBLrel. 01. Created)

DT 01-NOV-1996 (TrEMBLrel. 01. Last sequence update)
 DT 01-NOV-1996 (TrEMBLrel. 01. Last annotation update)
 DE HYPOTHEICAL 137.5 KDA PROTEIN (LPASP).
 GN LPA5W OR YP8132.01 OR YPL012W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Hall J., Ahmed A., Bussey H., Fortin N., Friesen J.D., Storms R.K.,
 RA Vo D.H., Wang Y., Winnett E.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 220-1228 FROM N.A.
 RA Badcock K., Churcher C.;
 RC STRAIN=AB972;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE OF 220-1228 FROM N.A.
 RA Walsh S.V., Barrell B., Rajandream M.A.;
 RC STRAIN=AB972;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE OF 220-1228 FROM N.A.
 RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
 RA Walsh S.V., Barrell B.G.;
 RC STRAIN=AB972;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U33335; AB68093.1; -;
 DR EMBL; 248483; CAAB8374.1; -;
 DR EMBL; 271255; CAAB95029.1; -;
 DR EMBL; 271255; CAAB95029.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 1228 AA; 137507 MW; 92333BCD53CB095F CRC64;

Query Match 4.7%; Score 111.5; DB 3; Length 1228;
 Best Local Similarity 19.9%; Pred. No. 4.4;
 Matches 111; Conservative 71; Mismatches 152; Indels 225; Gaps 28;
 Qy 11 KSLSETCPNGINGTKDARK-----VTGVIGSGDFAKSLTIRLI 50
 Db 222 KTMALTTISSGLAENKYLRLDTIFALPNSWVDTLLTNSWIAVINGMSTYATHQPKAL 351
 Qy 51 R-----CGYHVIGSRNPKF---ASEFPFHVVDVTHREDAL----- 83
 Db 352 RKIPGVFHMCTY---LASETEPVQAASQCLISILSESVKDDLLYTPSVDEKVKND 408
 Qy 84 -----TKNIIIFVAHREHYTSLMDRLHLVGLKILIDVSNMNRINQYPESNAEYLALF 137
 Db 409 ELISQIAKTFIDFLSIRYSHCS-----REIL--KILVAAFNKFY---RSNPFLKSL- 456
 Qy 138 PDSLIVKGFNVYSANALQGLPKDASROVYICSNINQAOQVIELAROLNFI----- 188
 Db 457 -----KIVDTWRVN-----EEQFMDLRNTEIELVIGASISANG 488
 Qy 189 -----PIDGSLSSAREIENPLRLFTLWRGPVW-----VALSLATFFLYSFVRDI 236
 Db 489 PEMILAEAPLNDNFSER-----FGR---AWLLPLIRDYTKNANLATP-----QNEL 533
 Qy 237 HPYARNOQSDFKIPIEIVN-----KTLP-----IVAITLSL 269
 Db 534 APYIKSFQSKFKVPEESIQLRVFOIVDQIWSLTPRCFCELPMDLRPSFTDEFASLSSL 593
 Qy 270 VYL-----AGLAAAYOLYYGTK-----YRRFP-----PWLEWLOCRK 303
 Db 594 LYSEVELRTTICHALKVLAESNVYAESSSHNVLLQRFPISEAKNTLEYLST-----K 648
 Qy 304 QUGLLSFFFAVHVAYSLCPMRS-----ERYL-----FLNMAYOOVHANISWNE 351.
 Db 649 STNLLAVLNV---YTQTPNARSYLLETIDQYLKITSKEDLEKTFNVVCGLLKSNWE 704
 Qy 352 EEWRIEYVIFGIMSLGLLSL--AVTSPISVSNALNWRFSFIQSTGLGVALLISTFH 409

705 E SGNVNKKKKPQTATLLDLCITMITYLPVSSYSALFSMES-----LTNSAD 754
QY 410 VLIYGNKRAPEEEVYRYT 428
Db 755 ALIO--KRA-----YRIIT 766

Search completed: March 14, 2001, 16:13:41
Job time: 82 sec